Modelling time evolving interactions in networks through a non stationary extension of stochastic block models

Marco Corneli, Pierre Latouche, Fabrice Rossi

Modelling time evolving interactions in networks through a non stationary extension of stochastic block models

Marco Corneli, Pierre Latouche and Fabrice Rossi

Université Paris 1 Panthéon-Sorbonne - Laboratoire SAMM
90 rue de Tolbiac, F-75634 Paris Cedex 13 - France

Abstract

In this paper, we focus on the stochastic block model (SBM), a probabilistic tool describing interactions between nodes of a network using latent clusters. The SBM assumes that the network has a stationary structure, in which connections of time varying intensity are not taken into account. In other words, interactions between two groups are forced to have the same features during the whole observation time. To overcome this limitation, we propose a partition of the whole time horizon, in which interactions are observed, and develop a non stationary extension of the SBM, allowing to simultaneously cluster the nodes in a network along with fixed time intervals in which the interactions take place. The number of clusters (K for nodes, D for time intervals) as well as the class memberships are finally obtained through maximizing the complete-data integrated likelihood by means of a greedy search approach. After showing that the model works properly with simulated data, we focus on a real data set. We thus consider the three days *ACM Hypertext* conference held in Turin, June 29th - July 1st 2009. Proximity interactions between attendees during the first day are modelled and an interesting clustering of the daily hours is finally obtained, with times of social gathering (e.g. coffee breaks) recovered by the approach. We compare our results with those obtained through similar methods and discuss improvements obtained with respect to the classical SBM. Applications to large networks are limited due to the computational complexity of the greedy search which is dominated by the number K_{up} and D_{up} of clusters used in the initialization. Therefore, advanced clustering tools are considered to reduce the number of clusters expected in the data, making the greedy search applicable to large networks.

Keywords: Random graphs, time clustering, stochastic block models, integrated classification likelihood.